Supporting Information for "Transport upscaling in highly heterogeneous aquifers and the prediction of tracer dispersion at the Macrodispersion Experiment (MADE) site"

Marco Dentz[1], Alessandro Comolli[1,2],Vivien Hakoun[1,3], and Juan J. Hidalgo[1*]
[1] *Spanish National Research Council (IDAEA-CSIC), c/Jordi Girona 18, 08034 Barcelona, ES*
[2] *Université Libre de Bruxelles (ULB), Nonlinear Physical Chemistry Unit, CP 231, 1050, Bruxelles, Belgium and*
[3] *BRGM, University of Montpellier, Montpellier, 34000, France*

I. INTRODUCTION

This supporting information details the numerical flow simulation in Section II. The numerical implementation of the predictive upscaled random walk model is reported in Section III. The longitudinal concentration profiles for the MADE-1 experiments for a resolution of $\Delta x = 10^{-1}$ m, and the profiles in semi-logarithmic scale are displayed in Section IV. The profiles and peak concentrations as a function of time for different values of K_g are discussed in Section V.

II. FLOW SOLUTION

The hydraulic conductivity fields are generated using the spectral method of Ruan and McLaughlin [1] with an exponential covariance function. The flow problem (3) is solved numerically using a two-point flux finite volume method in 40 realizations of $f(\mathbf{x})$ whose spatial mean and variance are within a 0.05 tolerance interval around the target values of 0 and 5.9. The flow domains have the dimensions $20\ell_h$ in the direction of the mean gradient, $10\ell_h$ in the transverse horizontal direction, and $10\ell_v$ in vertical direction. The resolution is $2^8 \times 2^7 \times 2^5$ voxels.

III. NUMERICAL MODEL IMPLEMENTATION

In the following, we outline the basic steps of the numerical implementation of the stochastic time domain random walk model combined with mass transfer between mobile and immobile regions. First, we note that the steady state Eulerian velocity distribution $p_e(v)$ is given by the distribution show in Figure 1. The distribution $p_v(v)$ of particle speeds is obtained according to (2) by flux-weighting.

Step 1: Initialization Initial particle velocities $v(0)$ are sampled from $p_v(v)$, particle positions $x(0)$ are chosen from the initial particle distribution $\rho(x)$, which here is $\rho(x) = \delta(x)$.

Step 2: Propagation of normal scores The initial velocities $v(0)$ are then converted to the initial values $w(0)$ of the normal scores using the map (7). The normal scores $w(s)$ are propagated from their initial values $w(0)$ according to the Ornstein-Uhlenbeck process (8), which is discretized by using an Euler scheme. This gives

$$w(s + \Delta s) = w(s)\left(1 - \ell_c^{-1}\Delta s\right) + \sqrt{2\ell_c^{-1}\Delta s}\zeta(s), \tag{1}$$

where $\zeta(s)$ denotes a Gaussian random variable with zero mean and unit variance. The discretization Δs is chosen much smaller than ℓ_c. Here we set $\Delta s = 10^{-1}$ m.

Step 3: Propagation of particle position and time The particle positions are incremented at each random walk step by $\Delta s/\chi$, the particle times by $\Delta s/v_\ell(s)$,

$$x(s + \Delta s) = x(s) + \frac{\Delta s}{\chi}, \qquad t_a(s + \Delta s) = t_a(s) + \tau_a(s), \qquad \tau_a(s) = \frac{\Delta s}{v_\ell(s)}. \tag{2}$$

The current particle speed $v_\ell(s)$ is obtained according to Eq. (7).

Steps 2 and 3 are repeated until the maximum simulation time is reached. Particle positions are recorded at steps s if $t(s) \leq t_i$ and $t(s + \Delta s) > t_i$ with t_i the desired observation times.

* E-mail: marco.dentz@csic.es

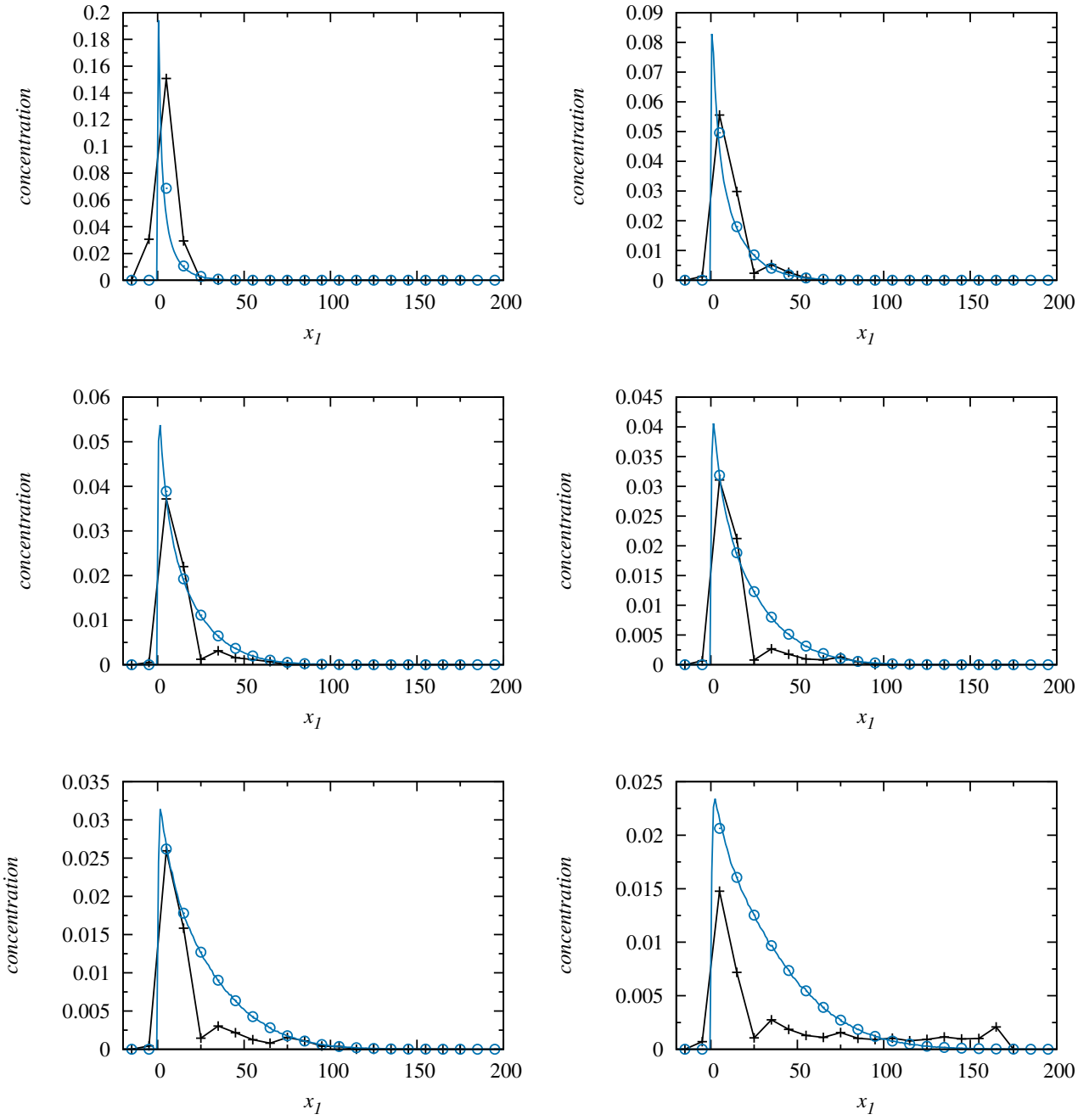

Figure 1. Concentration profiles from (crosses) the MADE data and the upscaled model for an averaging window of (circles) $\Delta x = 10$ m and (solid blue line) $\Delta x = 10^{-1}$ m at times (top left to bottom right) $t = 49$ d, 126 d, 202 d, 279 d, 370 d and 503 d.

IV. LONGITUDINAL CONCENTRATION PROFILES,

Figure 1 shows the longitudinal concentration profiles from the MADE-1 data, and the model prediction in coarse and fine resolution for $\theta = 1.2$. The comparison shows that the averaging window of $\Delta x = 10$ m introduces an artificial broadening of the peak due to oversmoothing, while the tail concentrations are nearly identical. With increasing time, as the peak width increases due to hydrodynamic dispersion, the differences between the coarse and fine scale models decrease.

Figure 2 shows the same data as in Figure 1 plotted logarithmically. The tail contains in general more mass than

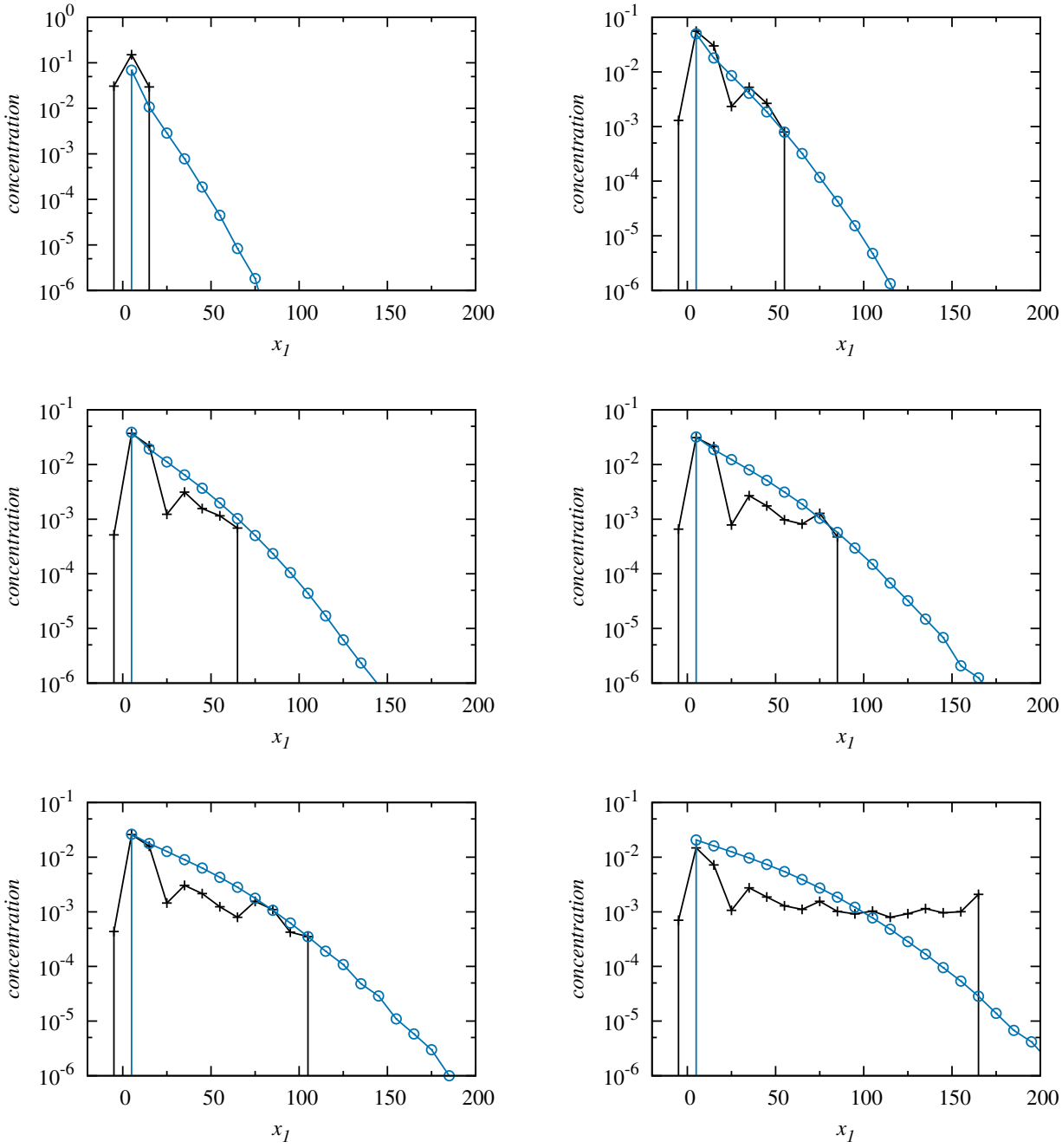

Figure 2. Same concentration data as in Figure 1 plotted logarithmically on the concentration axis.

the experimental data, but reproduces relatively well the overall tendency except for the 503 d snapshot at which mass is recovered at larger downstream distances than predicted by the model.

V. PREDICTIONS FOR DIFFERENT VALUES OF K_g AND PEAK CONCENTRATIONS

Figure 3 compares the model predictions for the MADE-1 data with $\theta = 1.2$ based on the estimated average value of $K_g = 4.3 \times 10^{-6}$ m/s, and the values at the upper and lower limits of the 95% confidence interval. $K_g = 6.7 \times 10^{-6}$ m/s and $K_g = 10^{-5}$ m/s. For all K_g values the model predicts the main plume features of peak localization and forward tail discussed in the main text. The model predictions based on $K_g = 5.5 \times 10^{-6}$ m/s shown in the main

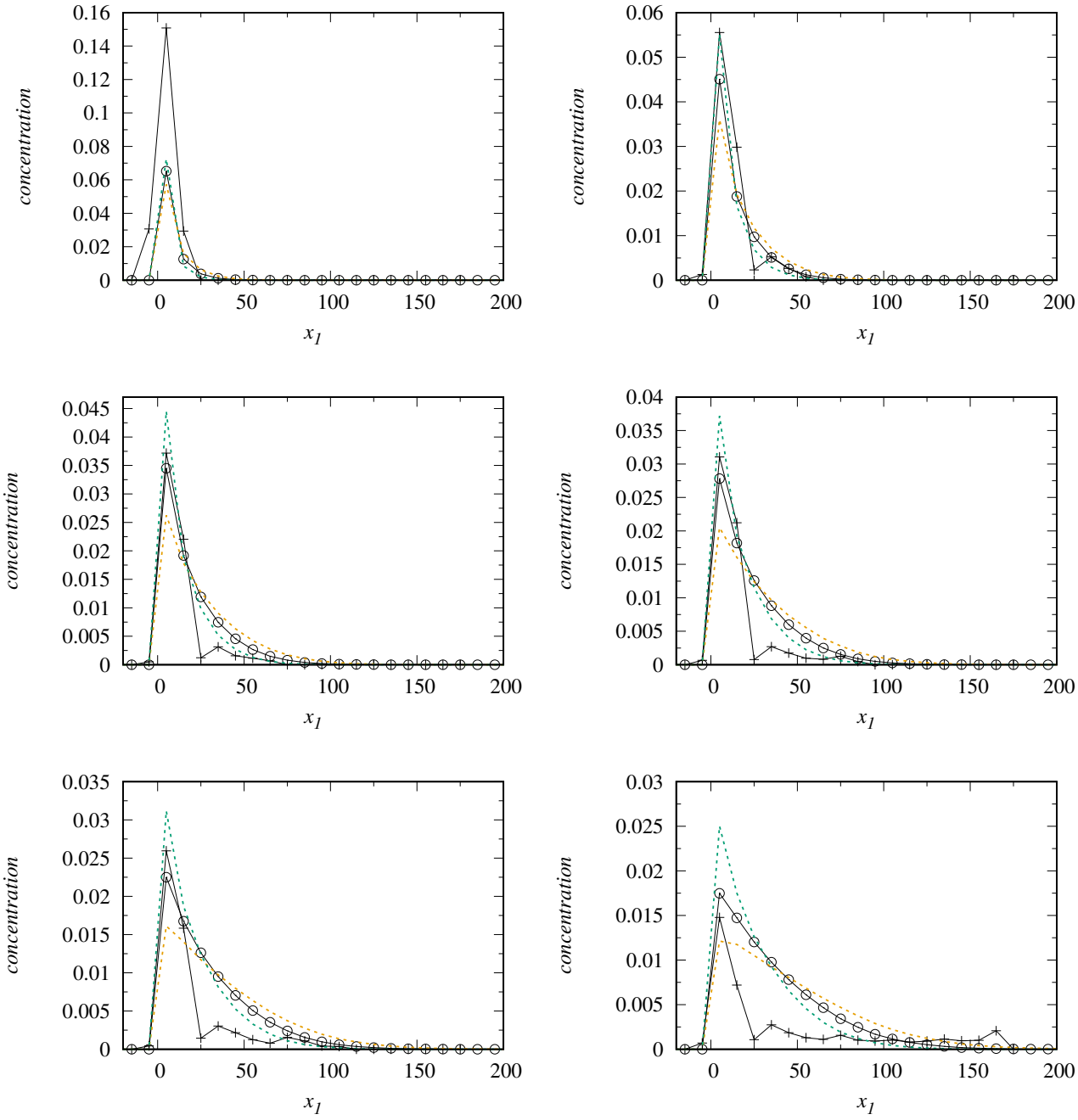

Figure 3. Concentration profiles from (crosses) the MADE-1 data and (circles) the upscaled model using $K_g = 6.7 \times 10^{-6}$ m/s for an averaging window of $\Delta x = 10$ m at times (top left to bottom right) $t = 49$ d, 126 d, 202 d, 279 d, 370 d and 503 d using a pointlike initial distribution. The dashed lines denote the predictions for (green) $K_g = 4.3 \times 10^{-6}$ m/s, (orange) $K_g = 10^{-5}$ m/s.

text provide slightly better estimates for the concentration peak values at $t \leq 370$ d than the ones using the mean $K_g = 6.7 \times 10^{-6}$ m/s.

The concentration peak is localized at all snapshots at $x = 5$ m, which indicates that it does not move beyond the averaging window between $x = 0$ m and 10 m. This is supported by the fine scale prediction. The upscaled model predicts peak heights and widths that are qualitatively and quantitatively similar to the experimental observations. The peak concentrations from the experimental data, and the corresponding model predictions are displayed in Figure 5. The upscaled model predicts the overall decay of the peak concentration. The decay behavior in the

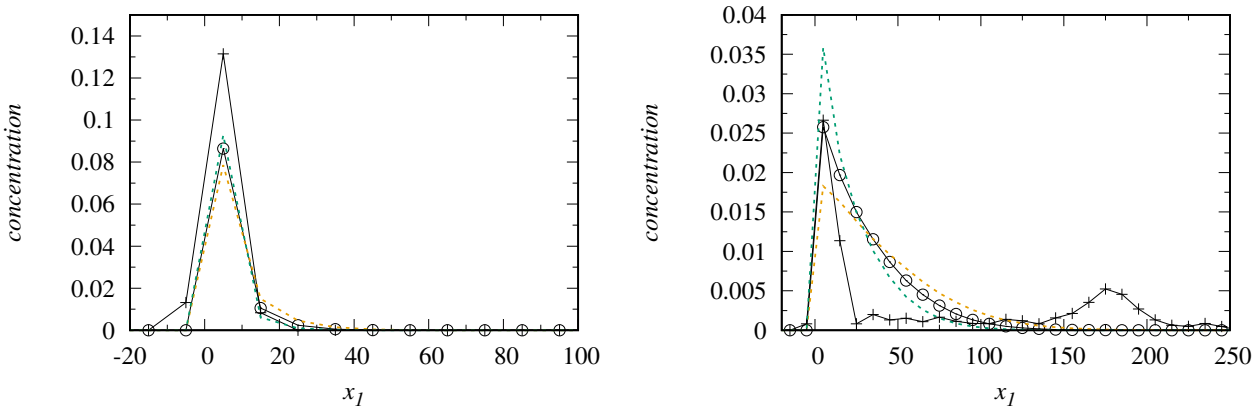

Figure 4. Analogous to Figure 3 for the MADE-2 data at $t = 27$ d, and 328 d.

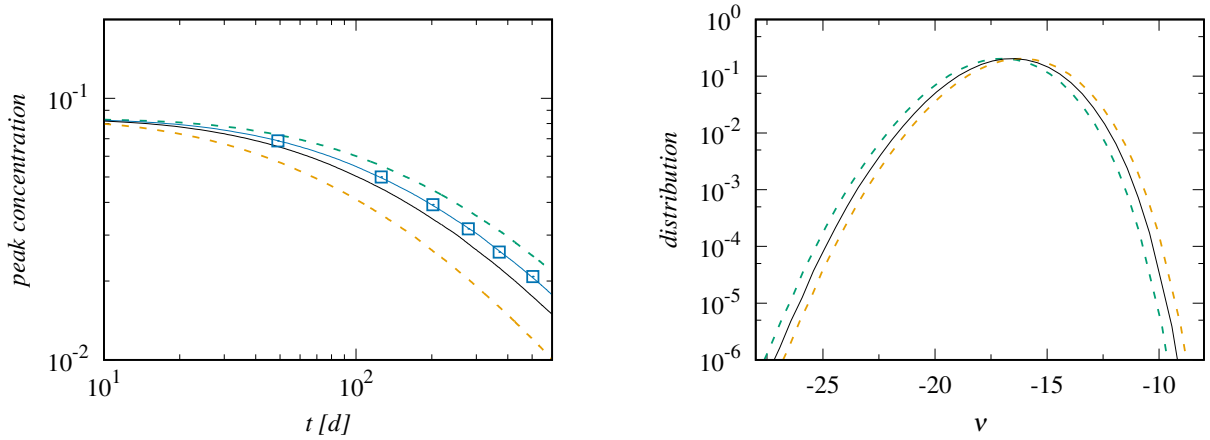

Figure 5. (Left panel) Maximum concentrations versus time from the (black crosses) MADE-1 data, and (blue squares) the prediction from the upscaled model with $\Delta x = 10$ m. The dashed lines denote the cumulative distribution of residence times (3) obtained from the upscaled model for (green) $K_g = 4.3 \times 10^{-6}$ m/s, (orange) $K_g = 10^{-5}$ m/s. The black solid line denotes the prediction for $K_g = 6.7 \times 10^{-6}$ m/s. (Right panel) Distributions of the corresponding $\nu = \ln(v)$.

localized peak is in fact a measure for the residence time in the injection region. The mass $m_0(t)$ remaining in the region between 0 m and 10 m is equal to the probability that the residence time τ_0 of a particle is larger than t,

$$m_0(t) = \int_t^\infty dt' \psi_0(t'), \tag{3}$$

where $\psi_0(t)$ denotes the residence time distribution. As long as the mass in the injection region of size Δx is larger than the mass in all other bins, the peak concentration is equal to $c_0(t) = m_0(t)/\Delta x$. Figure 5 shows the evolution of $c_0(t)$ from experimental data and model predictions obtained from the upscaled model. In order to highlight the impact of the value of geometric conductivity, we also plot the decay of the maximum concentration for the values $K_g = 6.7 \times 10^{-6}$ m/s and the values at the upper and lower limits of the 95% confidence interval, $K_g = 10^{-5}$ m/s and $K_g = 4.3 \times 10^{-6}$ m/s, as well as the corresponding distributions of $\nu = \ln(v)$. The former underestimates the observed maximum concentrations because, as shown in the right panel of Figure 5 the distribution of ν is characterized by higher values than for the lower K_g. This feature together with the persistence represented by the correlation length ℓ_c of particle velocities leads to a relatively fast tracer release from the injection domain compared to the other values for K_g. The value of $K_g = 4.3 \times 10^{-6}$ m/s overestimates the maximum concentration because of strong tracer retention at the injection region due to low initial speeds. Nevertheless, the predictions using these K_g values provide robust prediction of the salient non-Gaussian behaviors of the tracer profiles.

[1] F. Ruan and D. McLaughlin, Advances in Water Resources **21**, 385 (1998).